CLAIMS

Having thus described our invention, what we claim as new and desire to secure by Letters Patent is as follows:

1. A method for resource allocation where only a marginal utility is known, comprising the steps of:

determining an initial step size;

evaluating step size effectiveness with only information on the marginal utility;

reducing the step size if necessary; and

deploying an effective step size for utility optimization.

2. The method of claim 1, wherein allocated resources are database memory components.

3. The method of claim 1, wherein allocated resources are computer memory resources.

4. The method of claim 1, wherein allocated resources are computer system resources and resources are allocated to achieve optimal service level objectives.

5. An apparatus for resource allocation comprising:

a step size determination engine;

an optimization module for evaluating step size effectiveness with only information on the marginal utility;

a constraint module for reducing the step size if necessary; and

an interface module for deploying an effective step size for utility optimization.

6. The apparatus of claim 5, wherein allocated resources are database memory components.

7. The apparatus of claim 5, wherein allocated resources are computer memory resources.

8. The apparatus of claim 5, wherein allocated resources are computer system resources and resources are allocated to achieve optimal service level objectives.

9. A method of maximizing or minimizing an objective function $f(x)$, subject to constraints on a vector x where each vector x denotes a particular allocation of resources and the constraints generally describe properties of the resources which must be satisfied, the method maximizing or minimizing the objective function $f(x)$ while satisfying the constraints on x without a knowledge of f, said method comprising the steps of:

starting from an initial allocation, calculating a marginal utility of said allocation;

calculating constraint functions of said allocation;

applying the calculated constraint function information and marginal utility information to obtain a next allocation;

repeating the steps of calculating a marginal utility, calculating constraint functions and applying the calculated constraint function information and marginal utility information until a stopping criteria is satisfied; and

returning a locally optimal allocation of resources.

10. The method of maximizing or minimizing an objective function $f(x)$ as recited in claim 9, wherein said marginal utility is the gradient of the function f.

11. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein only the gradient ∇f is known and there is no procedure to evaluate the objective function f.

12. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein only the gradient ∇f and the Hessian $\nabla^2 f$ are known and there is no procedure to evaluate the objective function f.

13. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein the objective function is a utility function and the method maximizes the utility function.

14. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is time saved.

15. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is utilization of computer processors.

16. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is a number of transactions processed.

17. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 13, wherein said utility function is utilization of computer memory.

18. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, wherein the objective function is a cost function and the method minimizes the cost function.

19. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is power consumption.

20. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is total disk input/output time.

21. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 18, wherein said cost function is average system response time.

22. The method of maximizing or minimizing the objective function f (x) as recited in claim 9, wherein a backtracking line search is implemented in which a step α is decreased until the following condition is satisfied:

$$(\nabla f(x+\alpha p) - c_1 \nabla f(x))^T p \le 0,$$

where p is a search direction, c_1 is a constant, and $(\nabla f(x+\alpha p) - c_1 \nabla f(x))^T$ is the transpose of $(\nabla f(x+\alpha p) - c_1 \nabla f(x))$.

23. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein the allocated resources are database memory components.

24. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein allocated resources are computer memory resources.

25. The method of maximizing or minimizing the objective function $f(x)$ as recited in claim 9, the method optimally allocating resources wherein allocated resources are computer system resources and resources are allocated to achieve optimal service level objectives